A self-learning algorithm for biased molecular dynamics

Tribello, G. A., Ceriotti, M., & Parrinello, M. (2010). A self-learning algorithm for biased molecular dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(41), 17509-17514. https://doi.org/10.1073/pnas.1011511107

Published in:
Proceedings of the National Academy of Sciences of the United States of America

Document Version:
Peer reviewed version

A self-learning algorithm for biased molecular dynamics

Gareth A. Tribello, Michele Ceriotti and Michele Parrinello *

*Computational Science, Department of Chemistry and Applied Biosciences, ETHZ Zurich USI-Campus, Via Giuseppe Buffi 13 C-6900 Lugano

Submitted to Proceedings of the National Academy of Sciences of the United States of America

A new self-learning algorithm for accelerated dynamics, reconnaissance metadynamics, is proposed that is able to work with a very large number of collective coordinates. Acceleration of the dynamics is achieved by constructing a bias potential in terms of a patchwork of one-dimensional, locally valid collective coordinates. These collective coordinates are obtained from trajectory analyses so that they adapt to any new features encountered during the simulation. We show how this methodology can be used to enhance sampling in real chemical systems citing examples both from the physics of clusters and from the biological sciences.

accelerated dynamics | proteins | metadynamics | self-learning | sampling

Introduction

Many chemical systems, notably those in condensed matter and in biology, are characterized by the presence of multiple low energy states which are separated by large barriers. The presence of these barriers prevents exploration of all of configuration space during the relatively-short timescales accessible in molecular dynamics (MD) simulations. Typically this means that only those configurations in a small, locally-ergodic region in the vicinity of the input structure are visited.

A large number of methods have been put forward to overcome this difficulty, many of which either use some form of enhanced sampling (1–6) or focus on the transition from one local minimum to another (7–10). Other methods recognize that a small number of degrees of freedom (collective variables) accurately describe the interesting transitions and so either raise the temperature of these degrees of freedom (11–13) or introduce a bias to enhance the sampling along them (14–22). The major differences between the various approaches in this last class is the way in which the bias is generated a particularly useful technique being to use a bias that is dependent on the previously visited trajectory. This is the basis of the metadynamics method (23,24) that has been introduced by our group and applied to a large variety of chemical problems (25).

For many chemical systems the interesting, reactive processes take place in a relatively low dimensional space (26,27). However, it is often not immediately obvious how to identify a set of collective variables (CVs) that span this 'reactive' sub-space. Furthermore, in methods like metadynamics the presence of barriers in the transverse degrees of freedom leads to incomplete sampling. The obvious solution therefore is to use very large numbers of collective coordinates. However, although this is theoretically possible with methods such as metadynamics, it is impractical because the volume of bias that one must add to the free energy surface, and hence the length of the simulation, increases exponentially with dimensionality. One suggestion for resolving this issue is to run multiple short, 1D metadynamics simulations in parallel with different collective coordinates and to allow swaps between the different realizations based on a Monte Carlo criterion (28). Here we propose an alternative solution based on the realization that, if the free energy surface is to be flattened, the majority of the bias will have to be added at or near the minima in the surface. Identifying the locations of these minima is straightforward as, during a dynamical trajectory, the system should spend the majority of its time trapped in the vicinity of one or more of them (29). Therefore by using a form of "smart" bias that targets these low free energy regions specifically we can force the system away from them and into unexplored areas of configuration space. We call the self learning algorithm that we have developed based on these ideas Reconnaissance Metadynamics and have implemented it in the plugin for molecular dynamics PLUMED (24). In what follows we demonstrate this algorithm on two different systems - a cluster of seven Lennard Jones atoms and a short protein.

Background

Before introducing our new method a brief survey of established techniques for dealing with complex energy surfaces in terms of very large numbers of collective coordinates is in order. Zhu *et al* (30, 31) have introduced a method that uses a variable transformation to reduce barriers and thereby increase sampling, which works with a large number of collective coordinates. Problematically though this method does not work if there are hydrogen bonds present. In contrast, methods that work by thermostating the CVs at a higher temperature (11–13) suffer no such problems and have been used successfully with large numbers of CVs (32). However in these methods, unlike metadynamics, there is nothing that prevents the system from revisiting configurations, which could prove problematic for examinations of glassy landscapes with many local minima of equal likelihood.

For many free energy methods explicitly including a large numbers of collective coordinates is not feasible. However, one can use collective coordinates that describe a collective motion that involves many degrees of freedom. For example, one can use the principle components of the covariance matrix of a large set of collective coordinates, the values of which have been calculated over a short MD trajectory (33, 34). Alternatively, one can take non-linear combinations by defining a path in the high-dimensionality CV space (35). The distance along and the distance from this path span a low-dimensional, non-linear space and metadynamics simulations using these two CVs have been shown to work well. However, a great deal of insight is required in choosing an initial path from which the CVs are generated as the simulation can only pro-

Reserved for Publication Footnotes

vide meaningful insight for the region of configuration space in the immediate vicinity of this path.

Entropy plays an important role in free energy surfaces as it can wash out potential energy minima and make it such that finite temperature equilibrium states do not correspond to minima in the potential energy surface. Nevertheless, for systems with deep minima in the potential energy surface, one can use algorithms that locate all the minima on the 0 K (potential) energy surface and assume that the properties at every point on this surface are the same as those of the nearest local minima safe in the knowledge that the entropic effects are small. These algorithms allow one to divide up configuration space and map every point to its appropriate local minimum using a minimization algorithm (36). Furthermore, the slow modes in the vicinity of each basin provide good local collective coordinates as in the vicinity of basins the system is very nearly harmonic. From a patchwork of such descriptors one could conceive of ways to obtain globally-non-linear CVs. Recently, Kushima *et. al.* (37) have developed a self-learning, metadynamics-based algorithm based on these ideas that works by minimizing the energy and adding bias functions at the position of the minimum found so that subsequent minimizations will identify new minima.

Ideas described above for the exploration of potential energy surfaces cannot be straightforwardly transferred to the study of free energy surface because of the difficulties associated with the calculation of derivatives at finite temperature. Nonetheless, Maragakis *et al* (29) have developed a method, GAMUS, that has some similarities to the 0 K method developed by Kushima et al. In GAMUS the kinetic traps that are preventing free diffusion are located by fitting the probability distribution of visited configurations with a Gaussian Mixture (GM) model. The resulting set of bespoke Gaussians are then used to update an adaptive bias that encourages the system to visit unexplored regions of configuration space. This adaptive approach accelerates the filling of basins and thus provides a considerable speed up over conventional metadynamics when one is using 3 or 4 collective variables. Nevertheless, the filling time will still increase exponentially with the number of CVs.

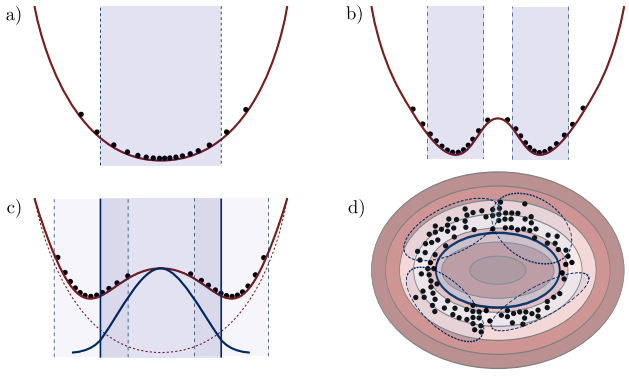

Fig. 2. Schematic representations of why it is necessary to use more than one pca analyzer and why it is necessary to expand basins. The trajectory in panel (a) can be fit using a single pca analyzer. By contrast the trajectory in panel (b) must be fit with a pair of analyzers as the energetic barrier is low enough that there will be hopping events between the two sub-basins. Panel (c) shows how adding a single Gaussian to the center of the basin in panel (a) can lead to the creation of spurious basins in later cluster analyses. Panel (d) demonstrates how this problem becomes more severe as the dimensionality is increased and also that it will occur if the basins are given a fixed size.

Reconnaissance metadynamics algorithm

Reconnaissance metadynamics combines a number of new ideas with those of established methods and is thus effective with very large numbers of CVs. The bias potential is constructed in terms of a patchwork of basins each of which corresponds to a low free energy region in the underlying FES. These features/basins are recognized dynamically by periodically analyzing the trajectory with a sophisticated clustering strategy. The region of configuration space in the vicinity of each basin is then described using a one-dimensional CV that is tuned using information collected during the clustering. Consequentially, even when the overall number of collective coordinates (d) is large, depressions are compensated for rapidly because the bias is added in a locally-valid, low-dimensional space.

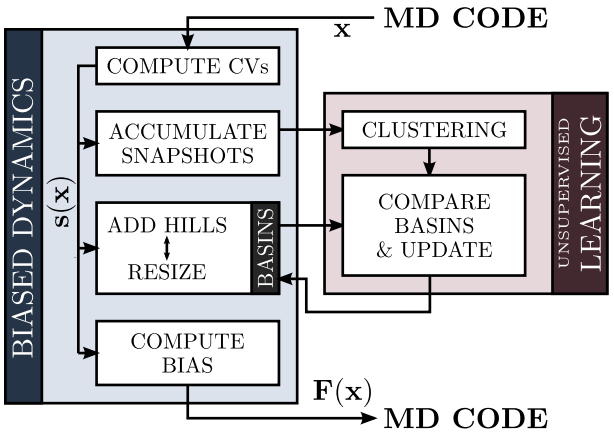

Fig. 1. Flow chart for the reconnaissance metadynamics algorithm.

Cluster analyses are performed at regular intervals using a set of stored configurations for the CVs that are accumulated from the trajectory. During this analysis it is essential that some form of dimensionality reduction be performed as otherwise the fitting will be intractable. In addition, one must recognize that, because this is a dynamical trajectory, the system may well have hopped between different basins on the free energy surface (see figure 1). Therefore, because we would ideally like to treat each basin separately, principle component analysis (PCA) is not an option. Furthermore, Gaussian mixture expectation maximization algorithms (29), although able to separate the basins, will become unstable when the number of collective coordinates is large. Thankfully however combinations of these two algorithms exist (38–40) that allow us to cluster the data while simultaneously reducing the dimensionality.

This clustering strategy provides us with a set of Gaussian centers ($\boldsymbol{\mu}$) and covariance matrices ($\mathbf{C}$) for the various basins in the free energy surface. Some of these will have very low weights or will be very similar to previously encountered basins and can thus be safely discarded. Those remaining provide useful information on the local topology of the FES but cannot be used to predict the actual depth of the basin or its shape away from the center. Consequentially, a flexible biassing strategy must be used in the vicinity of the minimum as addition of a single Gaussian will not necessarily compensate for the depression in free energy. This failure to compensate basins fully can lead to the formation of spurious low energy features in the region surrounding the basin center (see figure

1(c)), which is a problem that becomes more severe as the dimensionality is increased. To resolve these issues we assume that the basin is spherically symmetric in the metric induced by $\mathbf{C}$ and, in the spirit of metadynamics, construct an adaptive bias composed of small Gaussian hills of height w_i and width Δr, along a single, radial collective coordinate $r(\mathbf{s})$ (see equation 1).

$$V_i(\mathbf{s}) = w_i e^{-\frac{(r(\mathbf{s})-r_i)^2}{2\Delta r^2}} \quad [1]$$

$$r(\mathbf{s})^2 = (\mathbf{s}-\boldsymbol{\mu})^T \mathbf{C}^{-1} (\mathbf{s}-\boldsymbol{\mu}) \quad [2]$$

In the above $\mathbf{s}$ is a vector that denotes the position in the full, d-dimensional CV space. The form of equation 1 means that the bias associated with each hill acts in a hyper-ellipsoidal-crust-shaped region in the full, d-dimensional CV space. Each of these crusts has a shape much like a layer in an onion and so the integral of the bias added increases with $r(\mathbf{s})$. Consequentially, the filling time for each basin no longer depends exponentially on the number of collective coordinates. Furthermore, the hills have a shape that is consistent with the underlying anisotropy of the free energy surface because of the use of the covariance matrix in Eq. 2.

Obviously the distance from a basin center is only a good collective coordinate when we are close to that center and so each basin must have a size, S. This size assigns the region of configurational space in which it is reasonable to add hills that will force this system away from a particular basin. Setting an initial value for this size is straightforward as we know from our fitting that the basin's shape is well described by a multivariate Gaussian. Therefore, we choose an initial size equal to $S_0 = \sqrt{d-1}+3$ because, as shown in the supplementary information, when the angular dependency of a d-dimensional Gaussian is integrated out the resulting distribution of r is approximately equal to a 1D-Gaussian with standard deviation $\sqrt{2}$ centered at $\sqrt{d-1}$.

If the size of basins is fixed then the problems described in figure 1 are encountered once more. Hence, in reconnaissance metadynamics the basin's size is allowed to expand during the course of the simulation so as to ensure that spurious minima, that would otherwise appear at the edge of basins, are dealt with automatically. Periodically we check whether or not the system is inside the hyper-crust at a basin's rim ($S < r(\mathbf{s}) < S + \Delta r$) and also that it is not within the sphere of influence of any other basin. If these conditions are satisfied we then decide whether or not to expand using a probabilistic criterion, which ensures that it becomes more difficult to expand basins as the simulation progresses. Based on the loose analogy with free particle diffusion outlined in the supporting information, this probability for expansion is given by $P = \min(1, \frac{D\Delta t}{2\Delta r S})$, where S is the current size of the basin, Δt is the time between our checks on whether or not to expand and D is a user defined parameter.

The algorithm is summarized in the flow chart in figure . Further details on the components of the algorithm can be found in the methods section and in the supplementary information.

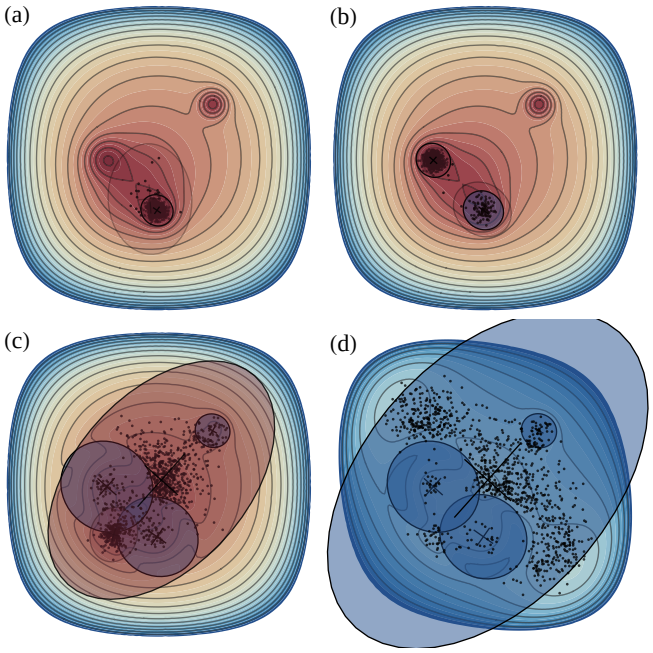

Fig. 3. Contour plots showing the potential energy + the current bias at selected points along a Reconnaissance Metadynamics trajectory for a particle diffusing about a 2D potential energy surface. The black dots indicate the positions of the snapshots accumulated from the trajectory while the red ellipses indicate the basins found using the PPCA algorithm. Blue ellipses are those basins, found during previous PPCA analyses, to which hills are being added. The expansion of these blue basins as the bias grows is clearly seen in this figure.

Results

2D-surface. To illustrate the operation of the algorithm we first show how it can be used to accelerate the (Langevin) dynamics on the model, 2D potential energy surface illustrated in figure .

At low temperatures a particle rolling about on the surface shown in figure will remain trapped in one of the deep basins. This is precisely what is observed during the first part of the simulation, when no metadynamics is performed, as the first application of our clustering algorithm demonstrates. On addition of bias, the system quickly escapes this first basin and falls into other basins, the locations of which are identified during subsequent applications of our unsupervised learning protocol. This process continues throughout the simulation so, once basins are identified, the history dependent bias compensates for them quickly and hence the system rapidly explores the entirety of the energy surface.

Figure shows that the reconnaissance metadynamics algorithm, when properly applied, finds only basins that correspond to the true features in the free energy surface. In addition figure (d) shows how effective the adaptive bias is in dealing with regions where small basins are encompassed in larger depressions. It clearly shows that initially small hills are used to deal with the sub-basins, while later much larger hills are used to compensate for the super-basin.

Lennard-Jones 7. Small clusters of rare-gas atoms have a remarkably complex behavior despite their rather limited number of degrees of freedom. A particularly well studied example is the two-dimensional, seven-atom, Lennard-Jones cluster (41,42) for which 4 minima and 19 saddle points in the potential energy surface have been identified (43) (see figure 4). At moderate temperatures ($k_B T = 0.1\epsilon$) this system spends the majority of its time oscillating around the minimum energy structure, in which one of the atoms is surrounded symmetrically by the six other atoms. Infrequently however the system will also undergo isomerizations in which the central atom of the hexagon is exchanged with one of the atoms on the surface (43).

To test the reconnaissance metadynamics algorithm we examined this system using the coordination numbers of all the atoms (seven collective coordinates). In doing this we ne-

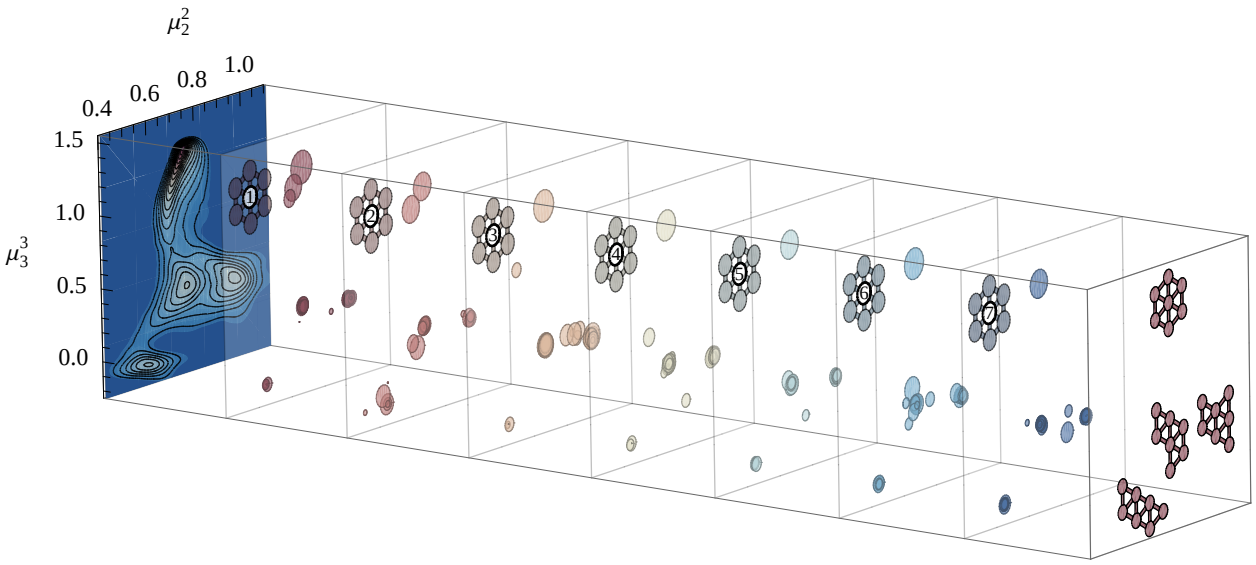

Fig. 4. The locations of the various basins found during a reconnaissance metadynamics simulation of the Lennard-Jones 7 cluster plotted as a function of the second and third moments of the distribution of coordination numbers and the index of the atom with the highest coordination number. Also shown is a free energy surface calculated using a well-tempered metadynamics simulation that employed the second and third moments of the distribution as collective coordinates. Each basin is represented by a circle with an area that is proportional to the bias at its center. On the free energy surface contours are placed at intervals of 0.1 ϵ

glect the interchange symmetry to see if we can reproduce this symmetry in the positions of the various basins we find. Figure 4 gives the results of this calculations and denotes the position of each of the basins found by a circle whose area reflects the final bias at the basin center. The three lowest lying minima for this cluster have one atom which has a much higher coordination number than the others and so the basins identified along the trajectory have been grouped, based on the index of the atom with the highest coordination number, onto seven slices. On each of these planes the positions and sizes of the basins are consistent, which suggests that the algorithm finds the correct symmetry and explores configuration space correctly.

In reconnaissance metadynamics there is no straightforward connection between the bias and the free energy because each CV is only valid in a local region and hence the overall bias is not a function of a global order parameter. Nonetheless, the bias greatly enhances the exploration of phase space and so free energies could be obtained using an umbrella sampling approach. However, even without this additional step, a reconnaissance metadynamics trajectory gives one a feel for the lie of the land, which can be used to obtain chemical insight. For example, the basin centers provide a set of landmark points that can be used to validate a low dimensionality description of the system (27). For this simple case we did this by calculating the value at the basin centers of many different candidate coordinates. We discovered that an optimal, in-plane separation of the various basins is attained when we use the second and third moments ($\mu_2^2 = \frac{1}{N}\sum_{i=1}^{N}(c_i - \langle c\rangle)^2$ and $\mu_3^3 = \frac{1}{N}\sum_{i=1}^{N}(c_i - \langle c\rangle)^3$ respectively) of the distribution of coordination numbers. These two CVs clearly project out permutation symmetry and so we also performed a conventional well-tempered metadynamics simulation. Figure 4 shows the free energy energy surface obtained in its eighth slice. A comparison of this surface with the results from the reconnaissance metadynamics shows how the basins found cluster around the minima in this FES.

Poly-alanine 12. The protein folding problem is commonly tackled using computer simulation and there exist model systems for which the entirety of the potential energy landscape has been mapped out (36) that represent a superb test of any new methodology. For example polyalanine-12, modeled with a distance dependent dielectric ($\epsilon_{ij} = r_{ij}$ in Angstroms) that mimics some of the solvent effects, has been extensively studied (44). This protein has a funnel-shaped, energy landscape with a alpha-helical, global minima. We found that during a 1 μs, conventional MD simulation started from a random configuration, the protein did not fold (see supplementary information). Hence, examining whether or not the protein will fold during a reconnaissance metadynamics simulation will provide a third test of our methodology.

For this reconnaissance metadynamics calculation we used the 24 backbone dihedral angles as the collective coordinates as these angles provide an excellent description of the protein structure. These variables are periodic, which had to be accounted for in the method by replacing the multivariate Gaussians with multivariate von Mises distributions (45). This distribution, if sufficiently concentrated about the mean, is equivalent to a Gaussian in which the difference between any point and the mean is shifted to the minimum image. Consequentially, we can continue to use the same algorithm for trajectory analysis as long as we take into account the periodicity when we calculate differences and averages. In addition, we can define a quantity (see equation 3) that is equivalent to the distance from the center of the basin (equation 2) but that takes into account the periodicity of the CVs (P_i).

$$\begin{aligned} r(\mathbf{s})^2 &= 2\sum_{i=1}^{d} C_{ii}^{-1}\left[1-\cos\left(\frac{2\pi[s_i-\mu_i]}{P_i}\right)\right] \\ &+ \sum_{i\neq j} C_{ij}^{-1}\sin\left(\frac{2\pi[s_i-\mu_i]}{P_i}\right)\sin\left(\frac{2\pi[s_j-\mu_j]}{P_j}\right) \end{aligned} \qquad [3]$$

Figure 5 provides a representation of a portion of a typical reconnaissance metadynamics trajectory of the protein. The figure shows the values of all the backbone torsional an-

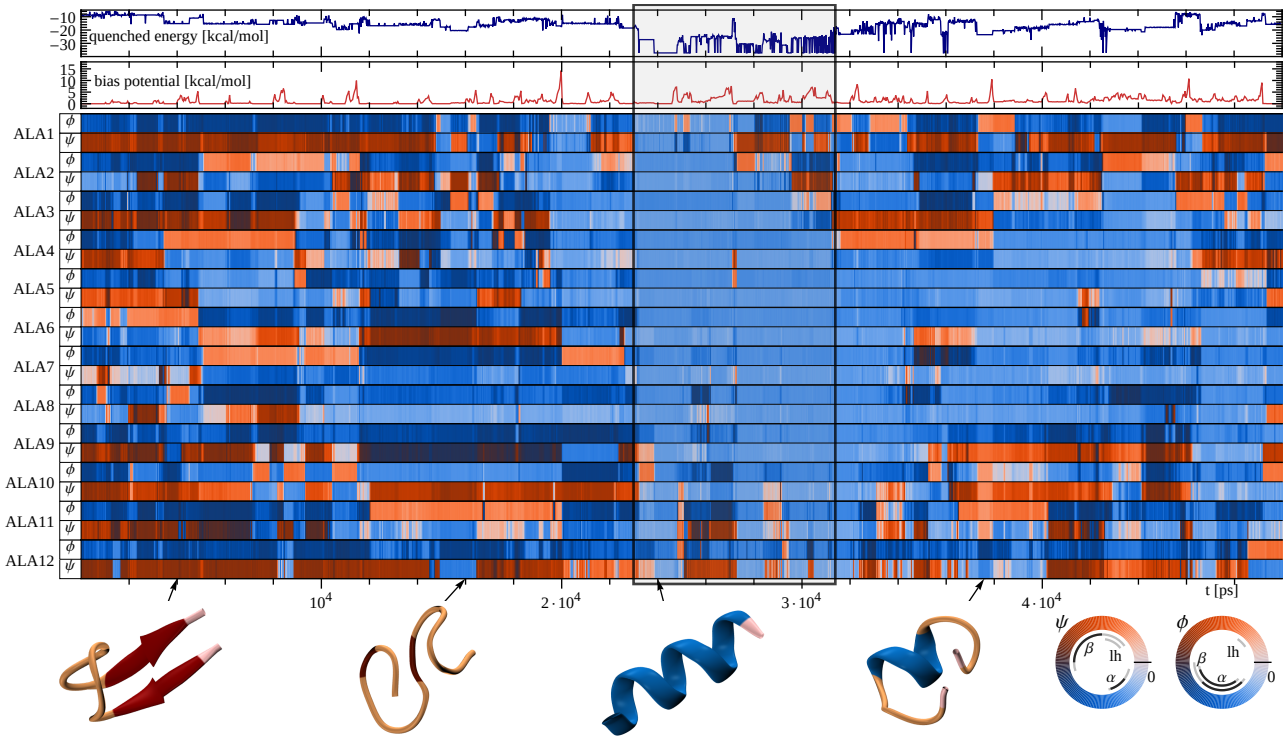

Fig. 5. A representation of a 50 ns portion of the reconnaissance metadynamics trajectory for alanine 12 modeled with a distance dependent dielectric. The bars in the lower panel give the values of 40 ps running averages for each of the torsional angles in the protein (see key). The red (light) line shows a similar running average for the bias potential. The system was annealed every 20 ps to examine the transitions between inherent structures and the energies obtained are indicated by the blue (dark) line. A number of low-energy, representative structures found during the trajectory are also shown and the box highlights the portion of the trajectory when the protein has a configuration that is near the global-minimum, alpha-helix structure.

gles and clearly demonstrates that a large volume of configurational space is explored during this relatively-short simulation. Furthermore, unlike in conventional MD, after approximately 22 ns the protein folds into the global-minimum, alpha-helix configuration. To further demonstrate that the reconnaissance metadynamics is ensuring that large portions of configurational space, that would otherwise not be visited, are being explored we calculated the inherent structures by minimizing the energy every 20 ps. The energies of the minimized structures are shown in the uppermost panel of figure 5 and demonstrate that the potential energy surface is very rough and that the alpha helix is considerably lower in energy than the other energy minima encountered.

Conclusions

We have proposed a new accelerated dynamics scheme, reconnaissance metadynamics, that uses a self-learning algorithm to construct a bias that accelerates the exploration of configuration space. This algorithm works by automatically identifying low free energy features and deploying a bias that efficiently and rapidly compensates for them. The great advantage of this algorithm is that there is no limit on the number of CVs on which the bias acts. This is only possible because we collapse all these CVs into a single collective coordinate that is only valid locally and patch together a number of these local descriptors in order to reflect the fact that the important degrees of freedom are not uniform throughout configuration space. As shown in figure this approach provides us with a simple, compact, hierarchical description of the free energy surface, which could be used to construct bias potentials for umbrella sampling.

As shown for the two chemical systems on which we have demonstrated the algorithm, our ability to work with large numbers of collective coordinates means that one can employ generic configurational data, such as torsional angles and coordination numbers, as collective coordinates and thereby avoid all the usual difficulties associated with choosing a small set of collective coordinates. In fact, even when CVs, on which there are no large barriers to motion, such as the torsional angles on the terminal amino acid groups in ala12, are included the algorithm will still function. For both systems our method produces trajectories that contain an extensive exploration of the low energy parts of configurational space and hence provides a feel for the lie of the land. Furthermore, even without a quantitative estimate of the free energy, considerable insight can be obtained from the trajectory. In the Lennard Jones this allowed us to attain an effective two dimensional description of the landscape. For more complex systems non-linear embedding could be used to automate this procedure.

Materials and Methods

Mixture of probabilistic principle component analyzers. Throughout this work we use the annealing strategy outlined in reference (40) and in the supplementary information to do clustering. This algorithm requires one to state at the outset the number of clusters that are being used to fit the data and the number of annealing steps. For the latter we initially set σ^2, which is the quantity treated like the temperature during the annealing, equal to the maximum eigenvalue of the covariance of the data and lowered σ^2 until it was less than 1 % of its initial value. To establish the correct number of clusters we run multiple fits to the data using different numbers of clusters and select the fit that gives the largest value for the Bayesian Informa-

tion Criterion (46), $\mathrm{BIC}{=}2\log[\mathcal{L}(\mathbf{x},\boldsymbol{\theta})]{-}n_p\log[M]$, where $\mathcal{L}(\mathbf{x},\boldsymbol{\theta})$ is the maximized likelihood for the model, n_p is the number of parameters in the model and M is the number of trajectory snapshots used in the fitting.

Selecting novel basins. As already discussed the GM algorithm provides us with the locations of a number of basins. Some of these will have very low weights in the fit and can therefore be safely ignored in the construction of the bias. Others however will provide information about the basins found during prior runs - in short information that is redundant. Therefore, we introduce a criterion for the selection of basins that requires that $f_i[1{-}\max(\xi_{ij})]{>}\mathrm{TOL}$, where TOL is some user defined tolerance, f_i is the weight of the new basin in the fit and ξ_{ij} is the similarity between the new basin i and the old basin j. To calculate ξ_{ij} we use Matusita's measure which can be calculated exactly for multivariate Gaussians (47). For an expanded basin of size S its original covariance is multiplied by a factor of S/S_0 when calculating this function so that their expanded volume is appropriately taken into account.

Lennard-Jones. The parameters for the simulations of Lennard-Jones 7, in Lennard Jones units are as follows. The temperature was set equal to $k_BT{=}0.1\epsilon$ using a Langevin thermostat, with a relaxation time of 0.1 $\sqrt{\epsilon/m\sigma^2}$. The equations of motion were integrated using the velocity verlet algorithm with a timestep of 0.01 $\sqrt{\epsilon/m\sigma^2}$ for 5 $\times$ 10^7 steps. During reconnaissance metadynamics the CVs were stored every 100 steps, while cluster analysis was done every 1 $\times$ 10^5 steps. Only basins with a weight greater than 0.3 were considered and to these attempts to add hills of height 0.5 k_BT and a width of 1.5 were made every 1000 steps. Basin expansion was attempted with the same frequency with the parameter D set equal to 0.03. The coordination numbers were computed using $c_i{=}\sum_{i\neq j}1{-}\left(\frac{r_{ij}}{1.5}\right)^8\left[1{-}\left(\frac{r_{ij}}{1.5}\right)^{16}\right]^{-1}$, where r_{ij} is the distance between atoms i and j.

poly-alanine-12. All simulations of polyalanine were run with a modified version of gromacs-4.0.3 (48), the amber96 forcefield (49) and a distance dependent dielectric. A timestep of 2 fs was used, all bonds were kept rigid using the LINCS algorithm and the van der Waals and electrostatic interactions were calculated without any cutoff. The global thermostat of Bussi et al (50) was used to maintain the system at a temperature of 300 K. The initial random configuration of the protein was generated by setting up the protein in a linear geometry, minimizing it and then running 1 ns of normal MD at 300 K in order to equilibrate. During reconnaissance metadynamics the CVs were stored every 250 steps, while cluster analysis was done every 5 $\times$ 10^5 steps. Only basins with a weight greater than 0.2 were considered and attempts were made every 1000 steps to add to these basins hills of height 0.4 k_BT and width 1.5. Basin expansion was attempted with the same frequency with the parameter D set equal to 0.3. Minimizations to obtain inherent structures was done by first annealing for 1.2 ns with a decay rate of 0.996 ps^{-1} and subsequently performing a steepest decent minimization. Guidelines as to how to select parameters for reconnaissance metadynamics are provided in the supplementary information. For both this system and the Lennard Jones cluster we obtained similar results with a variety of different parameters sets.

ACKNOWLEDGMENTS. The authors would like to thank Davide Branduardi, Jim Pfaendtner, Meher Prakash and Massimilliano Bonomi for useful discussions. David Wales is also thanked for his advice on the simulation of the ala12 system.

1. Sugita, Y, Okamoto, Y (1999) Replica-exchange molecular dynamics for protein folding. Chem Phys Lett. 314:141.
2. Hansmann, UE (1997) Parallel tempering algorithm for conformational studies of biological molecules. Chem Phys Lett. 281:140.
3. Marinari, E, Parisi, G (1992) Simulated tempering a new monte carlo scheme. Europhys. Lett. 19:451.
4. Liu, P, Kim, B, Friesner, RA, Berne, BJ (2005) Replica exchange with solute tempering: A method for sampling biological systems in explicit water. Proceedings of the National Academy of Sciences of the United States of America 102:13749–13754.
5. Nakajima, N, Higo, J, Kidera, A, Nakamura, H (1997) Flexible docking of a ligand and peptide to a receptor protein by multicanonical molecular dynamics simulation. Chemical Physics Letters 278:297 – 301.
6. Wang, F, Landau, DP (2001) Efficient, multiple-range random walk algorithm to calculate the density of states. Phys. Rev. Lett. 86:2050–2053.
7. Bolhuis, PG, Chandler, D, Dellago, C, Geissler, PL (2002) Transition path sampling: throwing ropers over rough mountain passes in the dark. Ann. Rev. Phys. Chem 54:20.
8. Faradjian, AK, Elber, R (2004) Computing time scales from reaction coordinates by milestoning. The Journal of Chemical Physics 120:10880–10889.
9. Allen, RJ, Warren, PB, ten Wolde, PR (2005) Sampling rare switching events in biochemical networks. Phys. Rev. Lett. 94:018104.
10. Maragliano, L, Fischer, A, Vanden-Eijnden, E, Ciccotti, G (2006) String method in collective variables: Minimum free energy paths and isocommittor surfaces. The Journal of Chemical Physics 125:024106.
11. Rosso, L, Minary, P, Zhu, Z, Tuckerman, ME (2002) On the use of the adiabatic molecular dynamics technique in the calculation of free energy profiles. J. Chem. Phys. 116:4961.
12. Maragliano, L, Vanden-Eijnden, E (2006) A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations. Chem. Phys. Lett. 426:168.
13. Abrams, JB, Tuckerman, ME (2008) Efficient and direct generation of multidimensional free energy surfaces via adiabatic dynamics without coordinate transformations. J. Phys. Chem. B 112:15742.
14. Patey, GN, Valleau, JP (1975) A monte carlo method for obtaining the interionic potential of mean force in ionic solution. The Journal of Chemical Physics 63:2334–2339.
15. Bash, P, Singh, U, Brown, F, Langridge, R, Kollman, P (1987) Calculation of the relative change in binding free energy of a protein-inhibitor complex. Science 235:574–576.
16. Ferrenberg, AM, Swendsen, RH (1988) New monte carlo technique for studying phase transitions. Phys. Rev. Lett. 61:2635.
17. Kumar, SK, Rosenberg, JM, Bouzida, D, Swendsen, RH, Kollman, PA (1995) Multidimensional free-energy calculations using the weighted histogram analysis method. J. Comput Chem 16:1339.
18. Huber, T, Torda, AE, van Gunsteren, WF (1994) Local elevation: a method for improving the searching properties of molecular dynamics simulation. J. Comput-Aided Mol Des 8:695.
19. Roux, B (1995) The calculation of the potential of mean force using computer simulations. Comput Phys Commun 91:275.
20. Voter, AF (1997) Hyperdynamics: Accelerated molecular dynamics of infrequent events. Phys. Rev. Lett. 78:3908.
21. Darve, E, Pohorille, A (2001) Calculating free energies using average force. The Journal of Chemical Physics 115:9169–9183.
22. Marsili, S, Barducci, A, Chelli, R, Procacci, P, Schettino, V (2006) Self-healing umbrella sampling: a non-equilibrium approach for quantitative free energy calculations. J. Phys. Chem. B. 110:14011.
23. Laio, A, Parrinello, M (2002) Escaping free energy minima. Proc. Natl. Acad Sci U.S.A. 99:12562.
24. Bonomi, M et al. (2009) Plumed: A portable plugin for free-energy calculations with molecular dynamics. Computer Physics Communications 180:1961 – 1972.
25. Laio, A, Gervasio, FL (2008) Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and materials sciences. Reports on Progress in Physics 71:126601.
26. Piana, S, Laio, A (2008) Advillin folding takes place on a hypersurface of small dimensionality. Physical Review Letters 101:208101.
27. Das, P, Moll, M, Stamati, H, Kavraki, LE, Clementi, C (2006) Low-dimensional, free-energy landscapes of protein-folding reactions by nonlinear dimensionality reduction. Proceedings of the National Academy of Sciences 103:9885–9890.
28. Piana, S, Laio, A (2007) A bias-exchange approach to protein folding. The Journal of Physical Chemistry B 111:4553–4559.
29. Maragakis, P, van der Vaart, A, Karplus, M (2009) Gaussian-mixture umbrella sampling. The Journal of Physical Chemistry B 113:4664–4673.
30. Zhu, Z, Tuckerman, ME, Samuelson, SO, Martyna, GJ (2002) Using novel variable transformations to enhance conformational sampling in molecular dynamics. Phys. Rev. Lett. 88:100201.
31. Minary, P, Tuckerman, ME, and Martiyna, GJ (2007) Dynamical spatial warping: A novel method for the conformational sampling of biophysical structure Siam. J. Sci. Comput. 30:2055
32. Abrams, CF, Vanden-Eijnden, E (2010) Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics. Proc. Natl. Acad. Sci. U.S.A. 107:4961.
33. Amadei, A, Linssen, ABM, Berendsen, HJC (1993) Essential dynamics of proteins. PROTEINS: struct. funct. gen. 17:412.
34. Grubmüller, H (1995) Predicting slow structural transitions in macromolecular systems: Conformational flooding. Phys. Rev. E 52:2893–2906.
35. Braduardi, D, Gervasio, FL, Parrinello, M (2007) From a to b in free energy space. J. Chem. Phys. 126:054103.
36. Wales, DJ (2003) Energy Landscapes (Cambridge University Press).
37. Kushima, A et al. (2009) Computing the viscosity of supercooled liquids. The Journal of Chemical Physics 130:224504.
38. Tipping, ME, Bishop, CM (1999) Probabilistic principal component analysis. J. R. Statist. Soc. B 61:611.
39. Tipping, ME, Bishop, CM (1999) Mixtures of probabilistic principal component analysers. Neural Computation 11:443.
40. Meinicke, P, Ritter, H (2001) Resolution-based complexity control for gaussian mixture models. Neural Computation 13:453.
41. Dellago, C, Bolhuis, PG, Csajka, FS, Chandler, D (1998) Transition path sampling and the calculation of rate constants. J. Chem. Phys. 108:1964.
42. Passerone, D, Parrinello, M (2001) Action-derived molecular dynamics in the study of rare events. Phys. Rev. Lett. 87:108302.
43. Wales, DJ (2002) Discrete path sampling. Molecular Physics 100:3285.
44. Mortenson, PN, Evans, DA, Wales, DJ (2002) Energy landscapes of model polyalanines. J. Chem. Phys. 117:1363.

45. Mardia, KV, Hughes, G, Taylor, CC, Singh, H (2005) A multivariate von mises distribution with applications to bioinformatics. Canadian Journal of Statistics 36:99.
46. Schwarz, G (1978) Estimating the dimension of a model. The Annals of Statistics 6:461.
47. Minami, M, Shmizu, K (1999) Estimate of similarity measure for multivariate normal distributions. Environmental and Ecological Statistics 6:229.
48. Hess, B, Kutzner, C, van der Spoel, D, Lindahl, E (2008) Gromacs 4: Algorithms for highly efficient, load-balanced and scalable molecular simulation. J. Chem. Theory Comput. 4:435.
49. Kollman, PA (1996) Advances and continuing challeges in achieving realistic and predictive simulations of the properties of organic and biological molecules. Acc. of Chem. Res. 29:461.
50. Bussi, G, Donadio, D, Parrinello, MJ (2007) Cannonical sampling through velocity rescaling. J. Chem. Phys. 126:014101.